



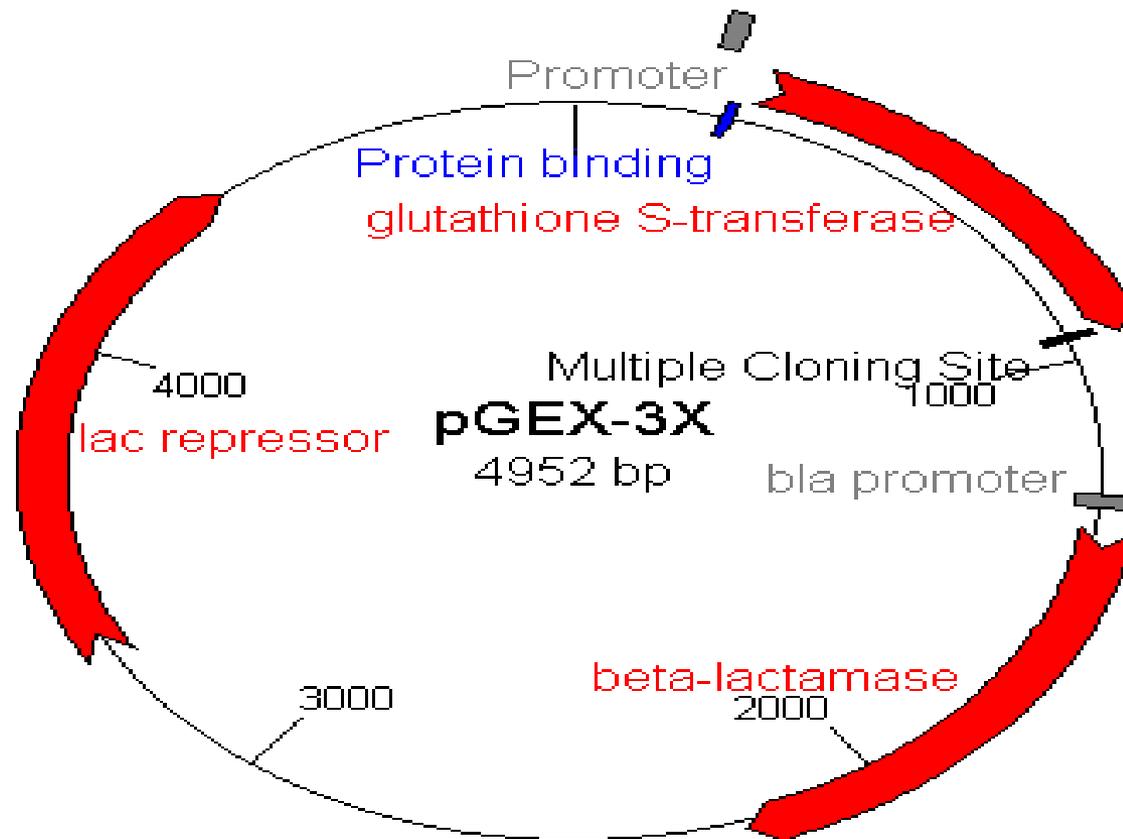
# Expression system

# expression vector

- ◆ Definition :

An **expression vector**, otherwise known as an **expression construct**, is usually a plasmid or virus designed for protein expression in cells

# pGEX-3X plasmid



# Difference between cloning vectors and expression vectors

- ◆ **Cloning vectors:**
- ◆ Cloning vectors are the DNA molecules that can carry a foreign dna segment in to the host cell. The vectors used in rDNA technology can be
- ◆ **Plasmids:** Self replicating, circular, extra chromosomal dna present in bacteria. plasmids have one or two copies per cell
- ◆ **Bacteriophages:** Virus infecting bacteria. bacteriophages have a high copy number per cell, so their copy number is also high in genome
- ◆ **Cosmids:** Hybrid vectors derived from plasmids which contain cos site of lambda phage

# Expression vectors

- ◆ The cloning vector containing suitable expression signals to have maximum gene for expression is called expression vector.
- ◆ The following expression signals are introduced in to gene cloned vectors to get maximum expression:
  - ◆ Insertion of a strong promoter.
  - ◆ Insertion of a strong termination codon.
  - ◆ Adjustment of distance between promoter and cloned gene.
  - ◆ Insertion of transcription termination sequence.
  - ◆ Insertion of a strong translation initiation sequence.

# Expression in Prokaryotes

**Promoter** – commonly used inducible promoters are promoters derived from *lac* operon and the T7 promoter.

Other strong promoters used include Trp promoter and Tac Promoter, which is a hybrid of both the Trp and Lac Operon promoters

**Ribosome binding site (RBS)** Follows the promoter, and promotes efficient translation of the protein of interest.

**Translation initiation site:** shine dalgarno sequence enclosed in the RBS, 8 base-pairs upstream of the AUG start codon

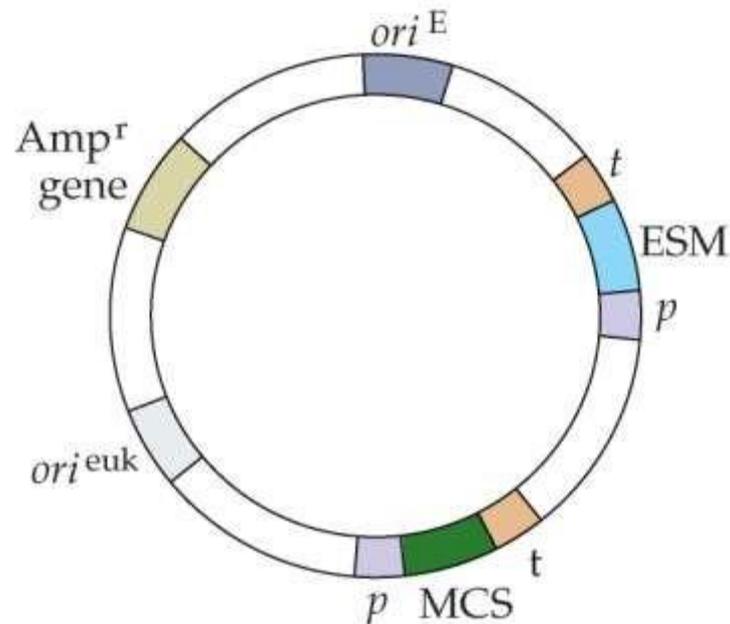
# Posttranslational Modification

- Eukaryotic protein synthesis occurs in cytoplasm or on the endoplasmic reticulum.
- These proteins are further post translational processed that is required for protein activity and stability.
- Disulfide isomerase also makes sure that the proteins produced have the correct configuration.
- The proper glycosylation that are necessary for protein conformation, localization by interacting with specific receptor and increase stability.

# Eukaryotic Expression Systems

- ◆ Eukaryotic expression vectors require sequences that encode for:
- ◆ **Polyadenylation tail**: Creates a polyadenylation tail at the end of the transcribed pre-mRNA that protects the mRNA from exonucleases and ensures transcriptional and translational termination: stabilizes mRNA production.
- ◆ **Minimal UTR length**: UTRs contain specific characteristics that may impede transcription or translation, and thus the shortest UTRs or none at all are encoded for in optimal expression vectors.
- ◆ **Kozak sequence**: Vectors should encode for a Kozak sequence in the mRNA, which assembles the ribosome for translation of the mRNA

- The major features of a eukaryotic expression vector are a promoter, a multiple cloning site, DNA segment for termination and polyadenylation, selectable marker, origin of replication in *E. coli* and eukaryotic cell and Amp<sup>r</sup> for marker in *E. coli*.



# Different types of expression system:

- ◆ **Prokaryotic expression systems**
- ◆ Escherichia coli
- ◆ Lactococcus lactis
- ◆ Other bacteria, e.g. Bacillus species

## yeast

- Pichia pastoris
- Pichia methanolica
- Saccharomyces cerevisiae

- ◆ **Insect cells**
- ◆ Baculovirus
- ◆ Stable recombinant cell lines
- ◆ Schneider cells (Drosophila)

## mammalian cells

- Viral infection,  
e.g adeno virus, retero virus ,lenti virus etc

- ◆ **In vitro expression systems**
- ◆ Rabbit reticulocyte lysate (red blood cells)
- ◆ Wheat germ extract
- ◆ Escherichia coli extract

## others

- Xenopus oocytes and cell-free extract
- Transgenic mice, plants
- Milk of transgenic animals

# Eukaryotic Expression Systems

- *Saccharomyces cerevisiae*
- *Pichia pastoris*
- Baculovirus-insect cell lines
- Mammalian systems

# *Saccharomyces cerevisiae*

- It is the most common eukaryotic system and there is a great deal of study about this organism.
- It is a single-celled and behaves like a bacterial culture and can be grown in relatively simple media in both small and large-scale production.
- Well characterized with many strong regulatable promoters with naturally occurring plasmids.
- Carry out post-translational modifications.
- Secretes very few of its own proteins.
- Recognized as safe by USDA and FDA.

**TABLE 7.1** Recombinant proteins produced by *S. cerevisiae* expression systems

---

Vaccines

Hepatitis B virus surface antigen  
Malaria circumsporozoite protein  
HIV-1 envelope protein

Diagnostics

Hepatitis C virus protein  
HIV-1 antigens

Human therapeutic agents

Epidermal growth factor  
Insulin  
Insulin-like growth factor  
Platelet-derived growth factor  
Proinsulin  
Fibroblast growth factor  
Granulocyte-macrophage colony-stimulating factor  
 $\alpha_1$ -Antitrypsin  
Blood coagulation factor XIIIa  
Hirudin  
Human growth factor  
Human serum albumin

---

HIV-1, human immunodeficiency virus type 1.

# *Saccharomyces cerevisiae*

- There are three main classes of *S. cerevisiae* expression vectors.
  - Yeast episomal plasmids (YEps).
  - Yeast integrating plasmids (YIps)
  - Yeast artificial chromosomes (YACs)
- Yeast episomal plasmids have been used extensively for the production of either intra- or extracellular heterologous proteins.
- Typically, vectors function in both *E. coli* and *S. cerevisiae*.

# *Saccharomyces cerevisiae*

- The YEps vectors are based on the high-copy-number 2 $\mu$ m plasmids.
  - The vectors replicate independently via a single origin of replication.
  - There are more than 30 copies per cell.
  - Selection scheme rely on mutant host strains that require a particular amino acid (histidine, tryptophan, or leucine) or nucleotide (uracil).
  - When a Yep with a wild-type *LEU2* gene is transformed into a mutant *leu2* host cell, only cells that carry plasmid will grow.

- Generally, tightly regulatable, inducible promoters are preferred for producing large amounts of recombinant protein at a specific time during large-scale growth.

**TABLE 7.2** Promoters for *S. cerevisiae* expression vectors

Promoter	Expression conditions	Status
Acid phosphatase ( <i>PH05</i> )	Phosphate-deficient medium	Inducible medium
Alcohol dehydrogenase I ( <i>ADHI</i> )	2–5% Glucose	Constitutive
Alcohol dehydrogenase II ( <i>ADHII</i> )	0.1–0.2% Glucose	Inducible
Cytochrome <i>c</i> <sub>1</sub> ( <i>CYC1</i> )	Glucose	Repressible
Gal-1-P Glc-1-P uridyltransferase	Galactose	Inducible
Galactokinase ( <i>GAL1</i> )	Galactose	Inducible
Glyceraldehyde-3-phosphate dehydrogenase ( <i>GAPD</i> , <i>GAPDH</i> )	2–5% Glucose	Constitutive
Metallothionein ( <i>CUP1</i> )	0.03–0.1 mM Copper	Inducible
Phosphoglycerate kinase ( <i>PGK</i> )	2–5% Glucose	Constitutive
Triosephosphate isomerase ( <i>TPI</i> )	2–5% Glucose	Constitutive
UDP-galactose epimerase ( <i>GAL10</i> )	Galactose	Inducible

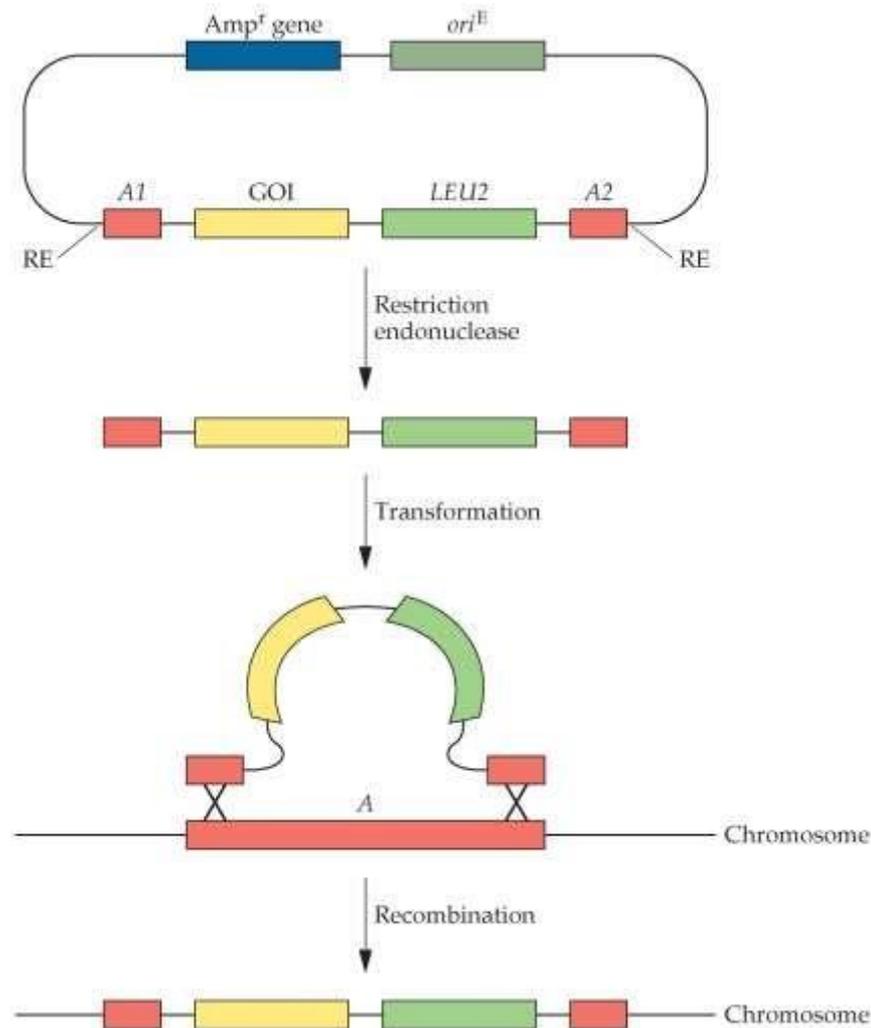
# *Saccharomyces cerevisiae*

- Most heterologous genes are provided with a DNA coding sequence for signal peptide that facilitates the secretion of protein through cell membranes and external environment.
- Other sequence that protect the recombinant protein from proteolytic degradation, and provide a affinity tag is also used.
- These extra amino acid sequences are equipped with a protease cleavage site so that they can be removed from the recombinant protein.

# *Saccharomyces cerevisiae*

- Plasmid-based yeast expression systems are often unstable under large-scale growth conditions even in the presence of selection pressure.
- A Yip vector is used to integrate a heterologous gene into the host genome to provide a more reliable production system.
- The plasmid does not usually carry an origin of replication.
- The disadvantage is the low yield of recombinant protein from a single gene copy.

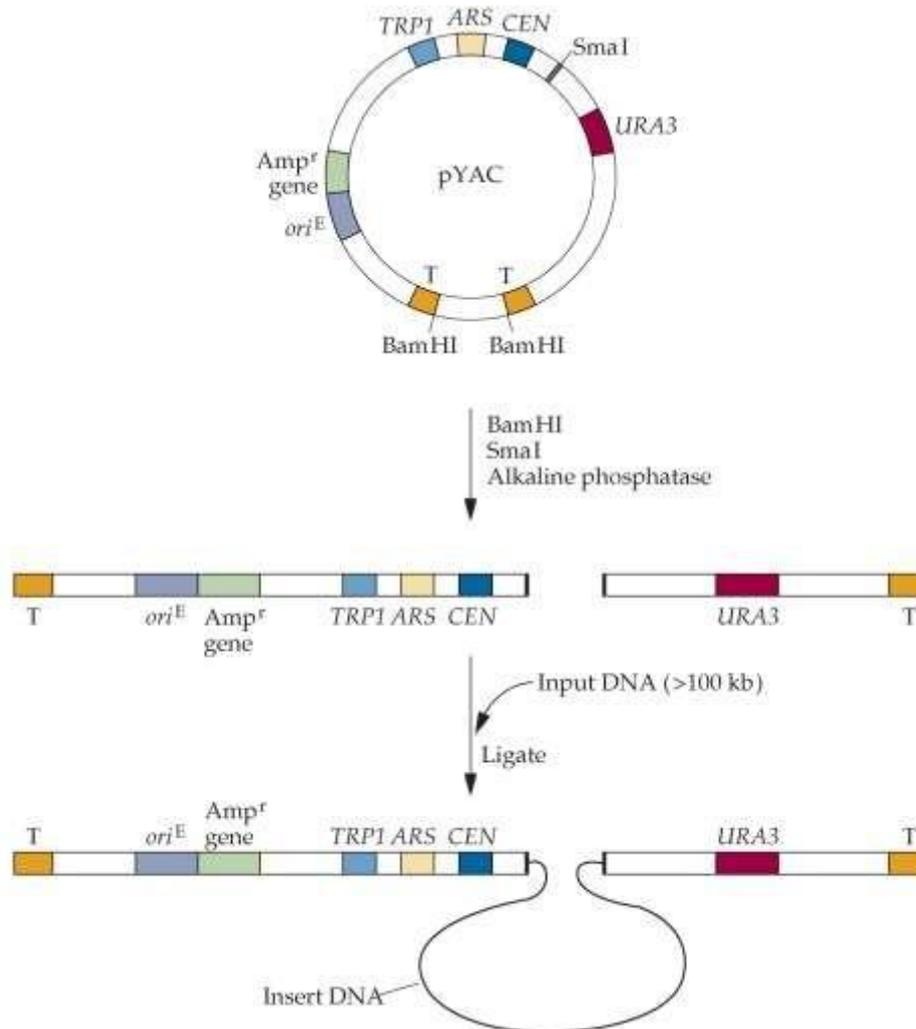
# Integration of DNA with a Yip vector



# YAC cloning system

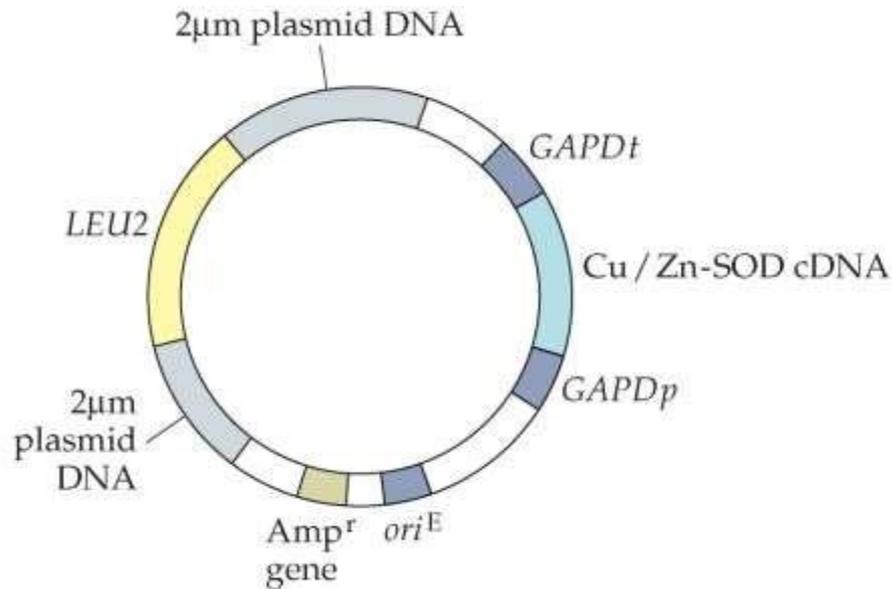
- A YAC is designed to clone a large segment of DNA (100 kb), which is then maintained as a separate chromosome in the host yeast cell.
- It is highly stable and has been used for the physical mapping of human genomic DNA, the analysis of transcription units, and genomic libraries.
- It has a sequences that act as ARS for replication, centromere for cell division, and telomere for stability.
- To date, they have not been used as expression systems for the commercial production.

# YAC cloning system



# Intercellular Production in Yeast

- Human Cu/Zn SOD cDNA was cloned between the promoter and termination-polyadenylation sequence of the yeast *GAPD* gene and subsequently used to transform LEU<sup>-</sup> mutant host cell.



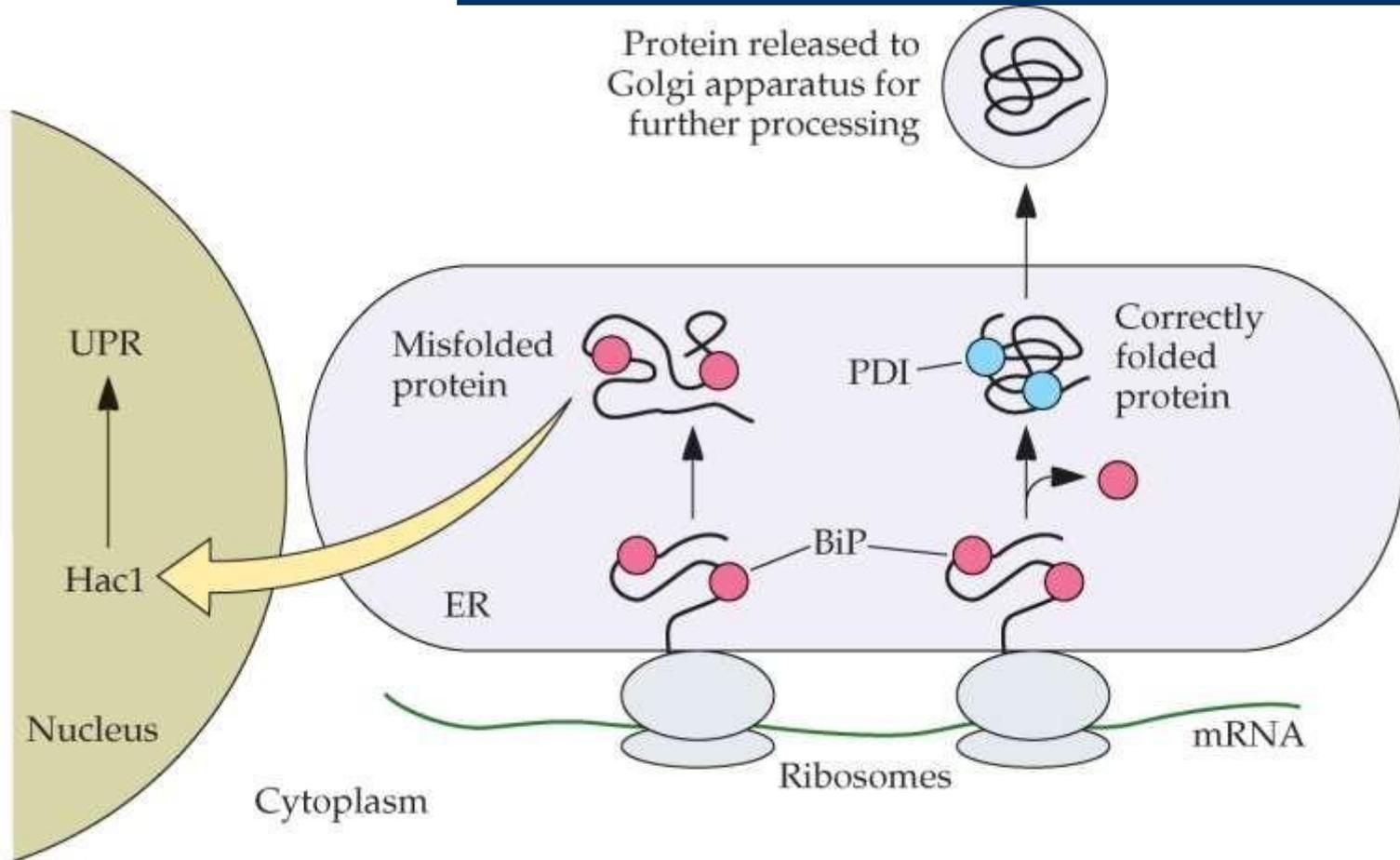
# Secretion of Heterologous Proteins

- Proteins may also be produced for secretion.
- In this system, any ~~glycosylated protein is secreted~~ (O or N linked).
- The coding sequences of recombinant proteins must be cloned downstream of a leader sequence, the yeast mating type factor  $\alpha$ -factor.
- Under these conditions, correct disulfide bond formation, proteolytic removal of the leader sequence, and appropriate posttranslational modifications occur, and an active recombinant protein is secreted.
- The leader peptide is removed by endoprotease that recognizes the Lys-Arg.

# Secretion of Heterologous Proteins

- For example, a properly processed and active form of the protein hirudin; a powerful anticoagulant protein cloned from a leech, was synthesized and secreted by an *S. cerevisiae*.
- A YEp vector that had the prepro- $\alpha$ -factor sequence added to the hirudin coding sequence to allow expression that is cleaved away in processing.
- Leaves active hirudin which is secreted.
- Producing a recombinant protein for use in human therapeutics in yeast rather than in bacteria is to ensure the proper folding.

# Secretion of Heterologous Proteins



# *Pichia pastoris* Expression Systems

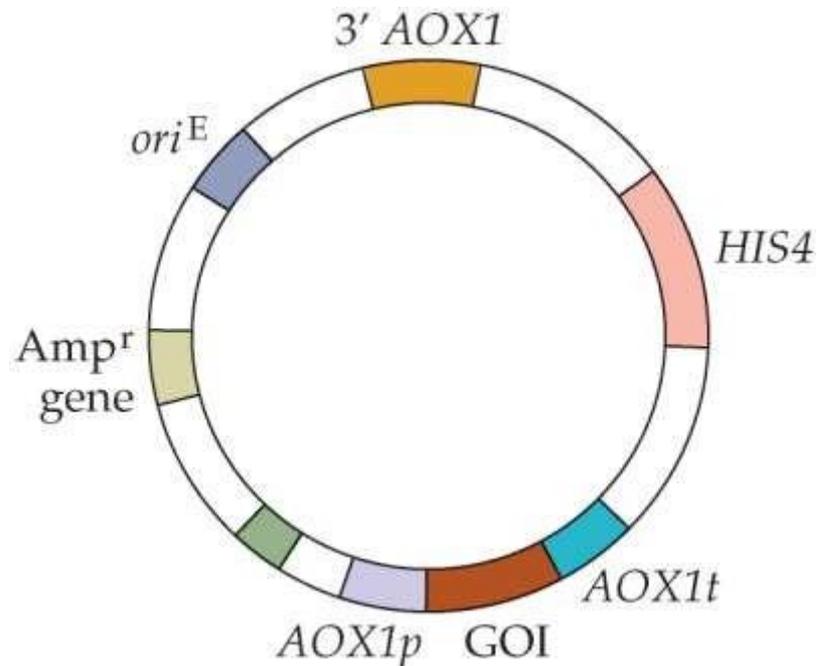
- Though *S. cerevisiae* is successfully used to produce recombinant proteins for human, it has major drawbacks.
- The level of protein production is low.
- There is the tendency for hyperglycosylation resulting in change of protein function.
- Proteins are often retained in periplasm, increasing time and cost for purification.
- It produces ethanol at high cell densities, which is toxic to cells.

# *Pichia pastoris* Expression Systems

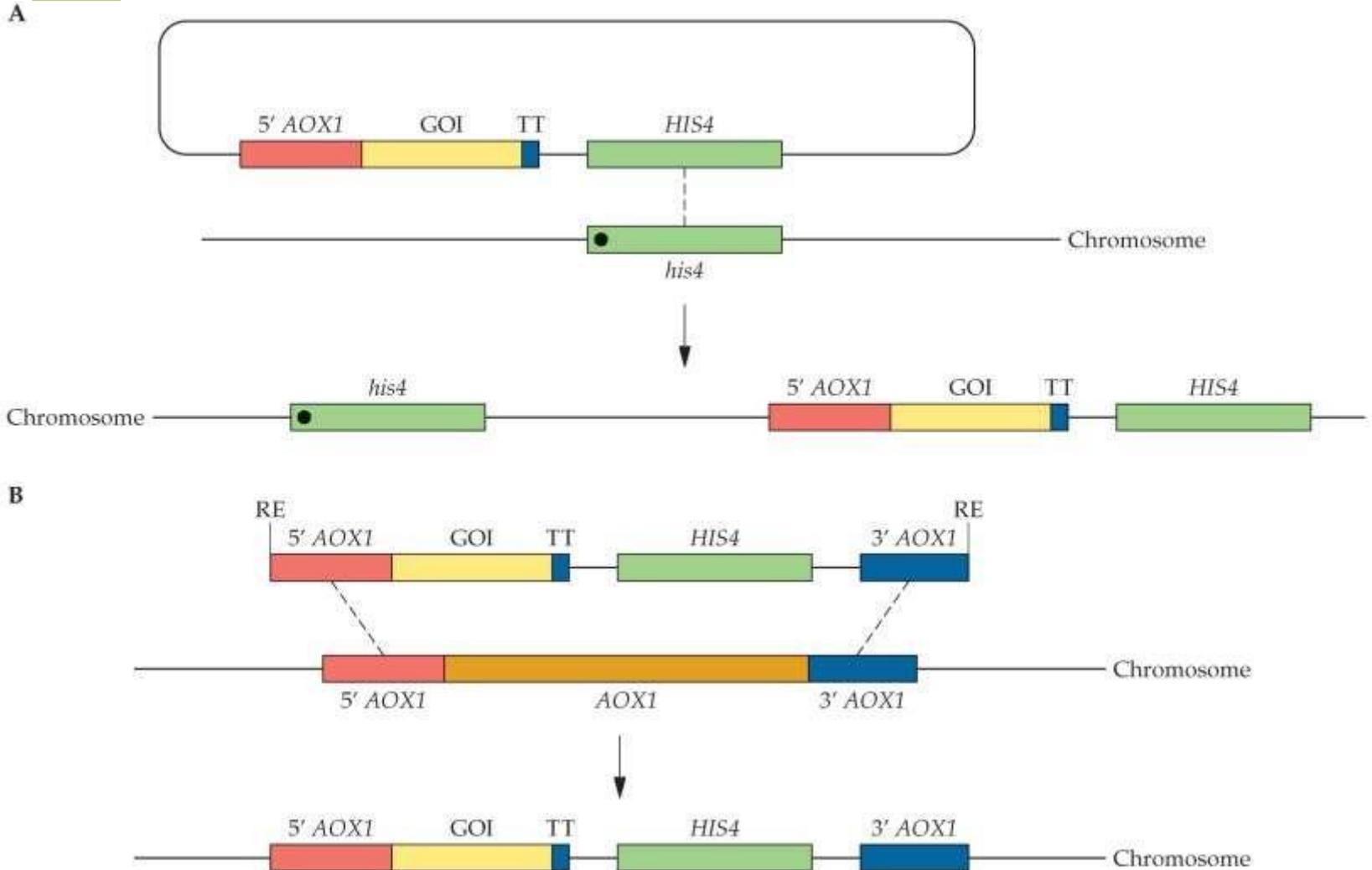
- *P. pastoris* is a methylotrophic yeast that is able to utilize methanol as a source of carbon and energy.
- Glycosylation occurs to a lesser extent and the linkages between sugar residues are of the  $\alpha$ -1,2 type.
- *P. pastoris* strain was extensively engineered with the aim of developing a “humanized” strain that glycosylate proteins in a manner identical to that of human cells.
- It does not produce ethanol.
- It normally secretes very few proteins, thus simplifying the purification of secreted recombinant proteins.

# *Pichia pastoris* Expression Systems

- A double recombination event between the *AOX1p* and *AOX1* regions of the vector and the homologous segments of chromosome DNA results in the insertion of the DNA carrying the gene of interest and the *HIS4* gene.



# *Pichia pastoris* Expression Systems



# Baculovirus-Insect Cell Expression

- Baculoviruses are a large, diverse group of viruses that specifically infect arthropods, and are not infectious to other animals.
- During the infection cycle, two forms of baculovirus are produced.
  - A single nucleocapsid (virus particle) which can infect more midgut cells.
  - Clusters of nucleocapsids that are produced outside of the cells (virions) in a protein matrix (polyhedrin).

# Baculovirus-Insect Cell Expression

- The polyhedrin gene is replaced with a coding sequence for a heterologous protein, followed by infection of cultured insect cells, resulting in the production of the heterologous protein.

**TABLE 7.6** Some of the recombinant proteins that have been produced by the baculovirus expression vector system

$\alpha$ -Interferon	G-protein-coupled receptors	Malaria proteins
Adenosine deaminase	HIV-1 envelope protein	Mouse monoclonal antibodies
Anthrax antigen	HSV capsid proteins	Multidrug transporter protein
$\beta$ -Amyloid precursor protein	Human alkaline phosphatase	Poliovirus proteins
$\beta$ -Interferon	Human DNA polymerase $\alpha$	Pseudorabies virus glycoprotein 50
Bovine rhodopsin	Human pancreatic lipase	Rabies virus glycoprotein
Cystic fibrosis transmembrane conductance regulator	Influenza virus hemagglutinin	Respiratory syncytial virus antigen
Dengue virus type 1 antigen	Interleukin-2	Simian rotavirus capsid antigen
Erythropoietin	Lassa virus protein	Tissue plasminogen activator

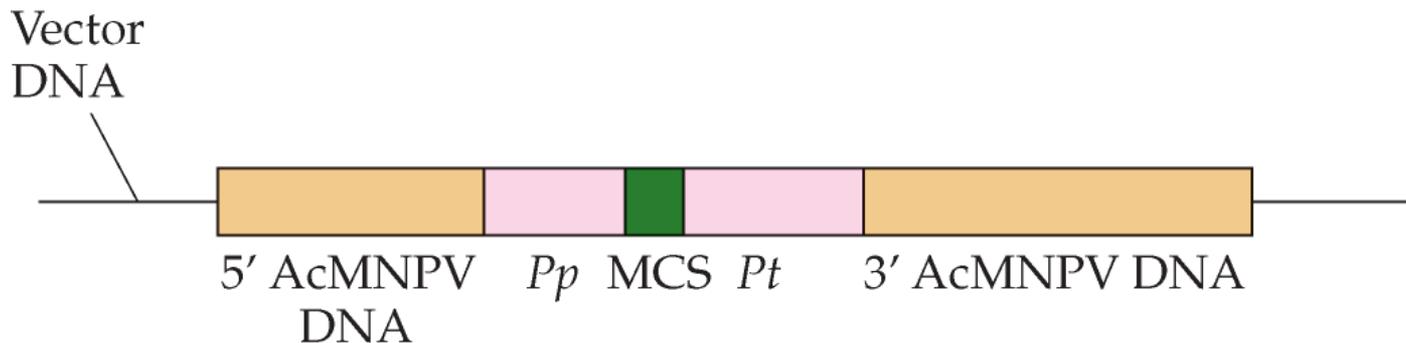
HIV-1, human immunodeficiency virus type 1; HSV, herpes simplex virus.

# Baculovirus-Insect Cell Expression

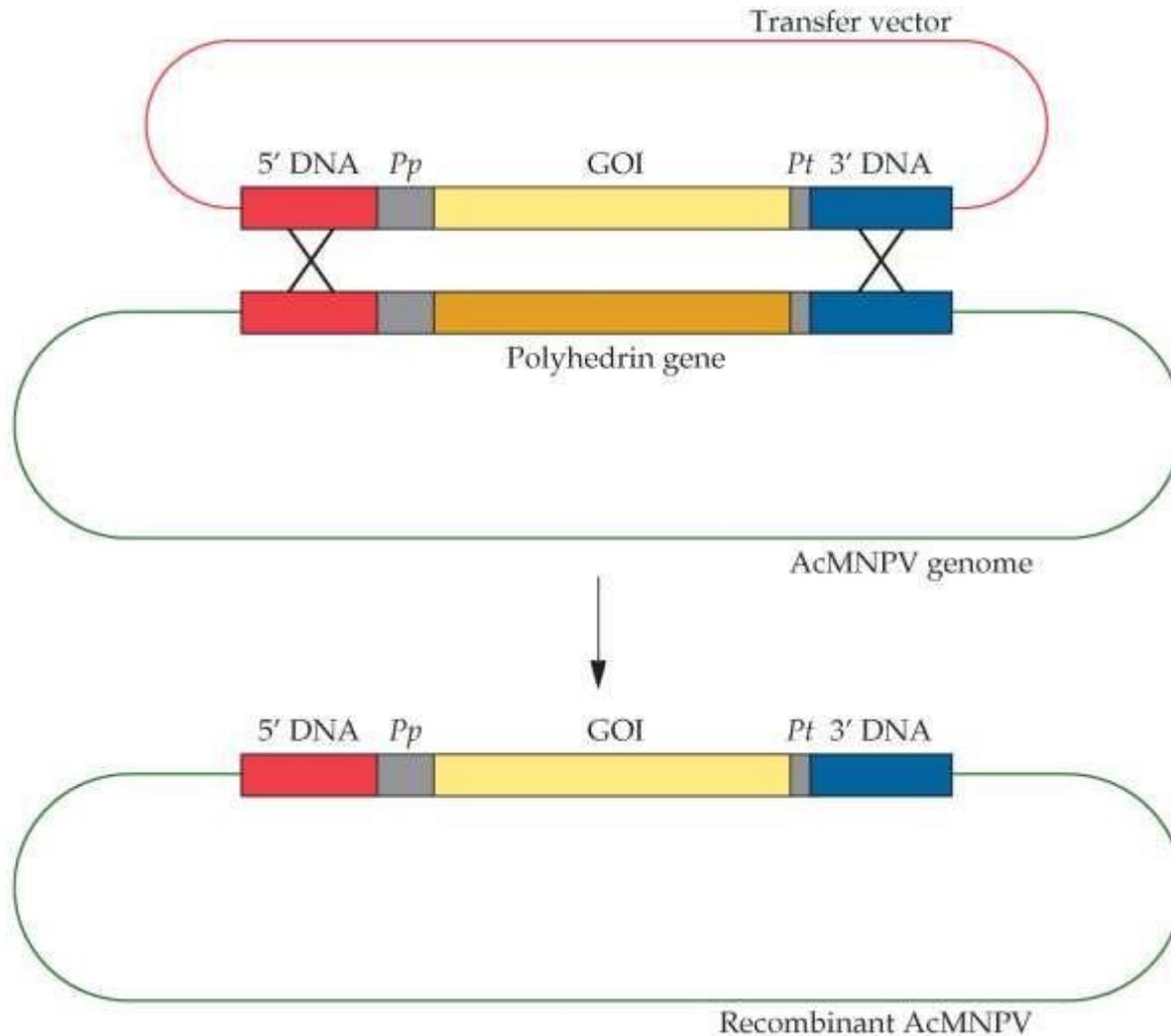
- Constructs have been made using the polyhedrin promoter to produce large quantities of extracellular protein.
- Most proteins are modified and secreted properly.
- Grows very well in many insect cell lines allowing easy production.
- Minor problem that doesn't process certain mammalian glycosylation types correctly (galactose and sialic acid; N-linked.)

# Baculovirus Expression Vectors

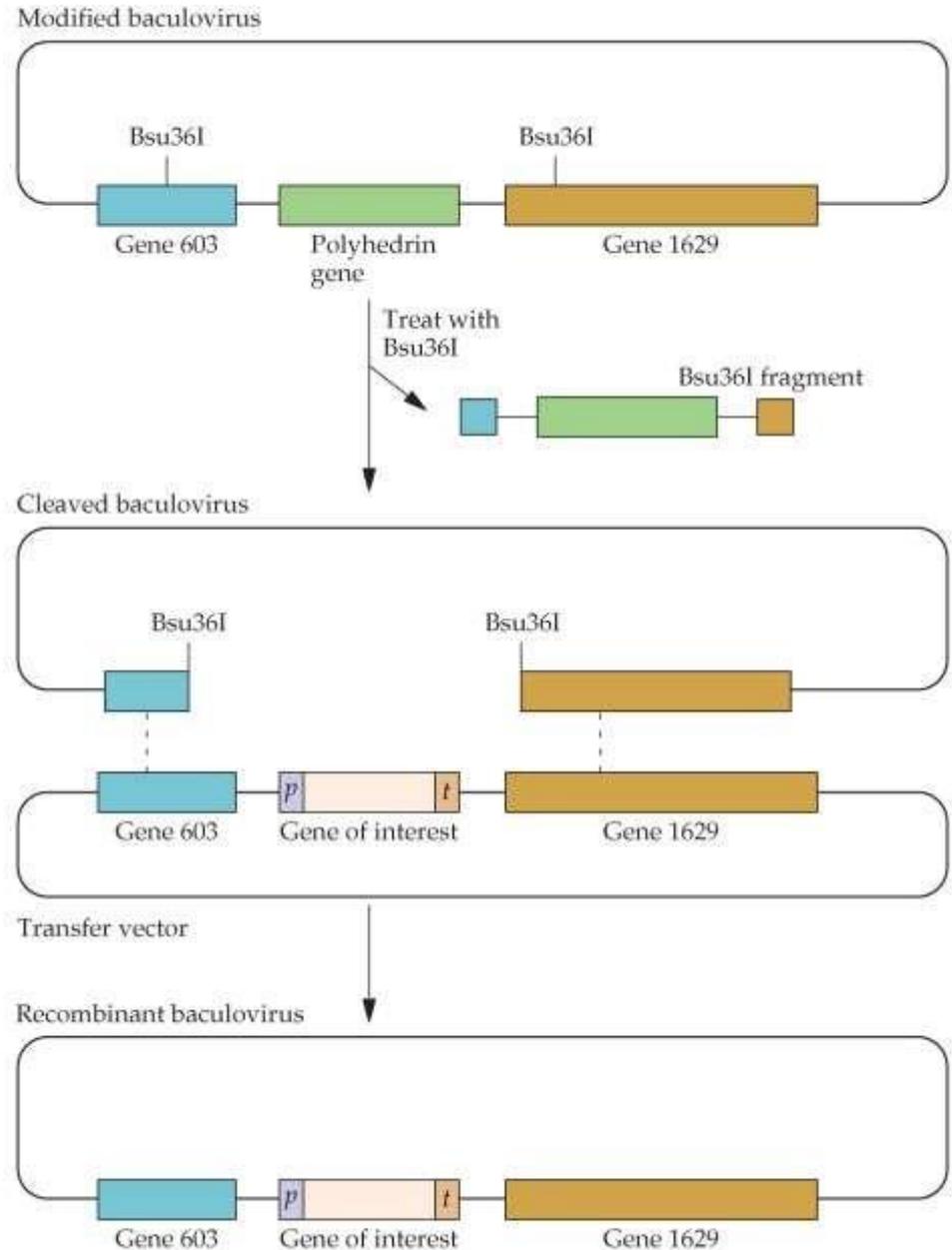
- The specific baculovirus that has been used extensively is *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV.)
- A gene of interest is inserted into the MCS and the transfer vector is propagated in *E. coli*.
- Next, insect cells in culture are cotransfected with AcMNPV DNA and the transfer vector carrying the cloned gene.



# Baculovirus Expression Vectors



# Increasing the Yield of Recombinant Baculovirus



# Mammalian Cell Expression Systems

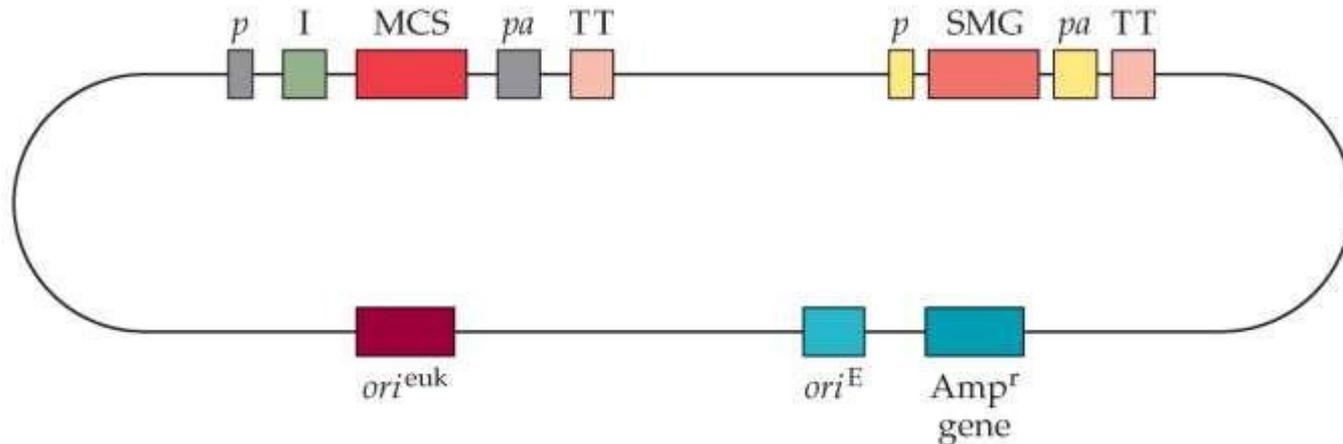
- Important for producing proteins with all post-translational modifications.
- Many established cell lines are useful.
- Transient expression: African green monkey, baby hamster, & human embryonic (all kidney tissue cell lines.)
- Long-term expression: Chinese hamster ovary and mouse myeloma cells.

# Mammalian Cell Expression Systems

- Expression vectors in these systems are usually derived from an animal virus such as SV40 (simian virus 40).
- Can be used for expression of single polypeptides, homooligomers, and heterooligomers.
- The latter is made possible by transforming with two or more separate cloned genes.
- Industrial production is however costly.

# Vector Design

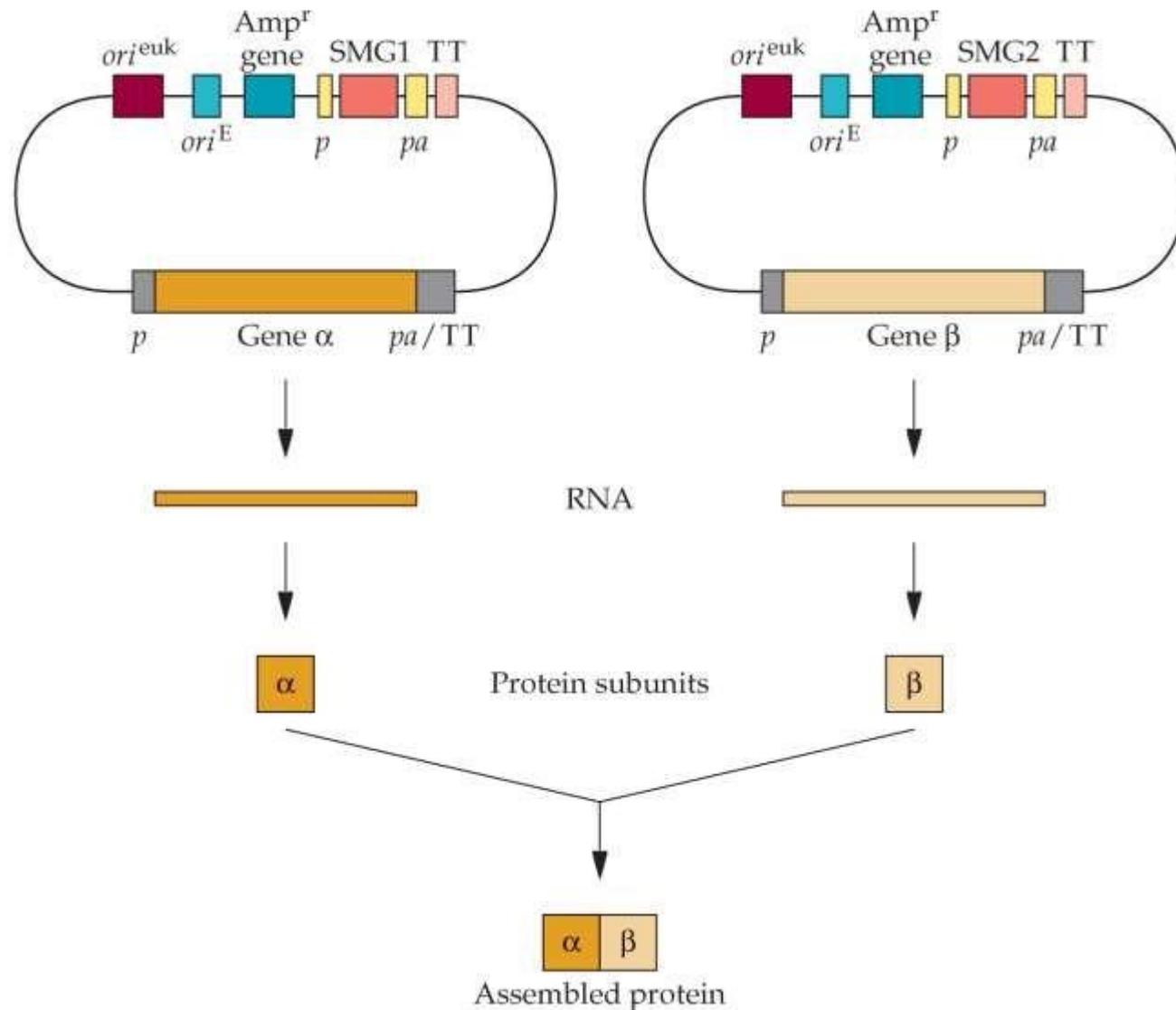
- Generalized mammalian expression vector.
- The MCS and SMG are under the control of eukaryotic promoter, polyadenylation, and terminal sequence.
- An intron enhances the production of heterologous protein.
- The Amp<sup>r</sup> gene is used for selecting transformed *E. coli*.



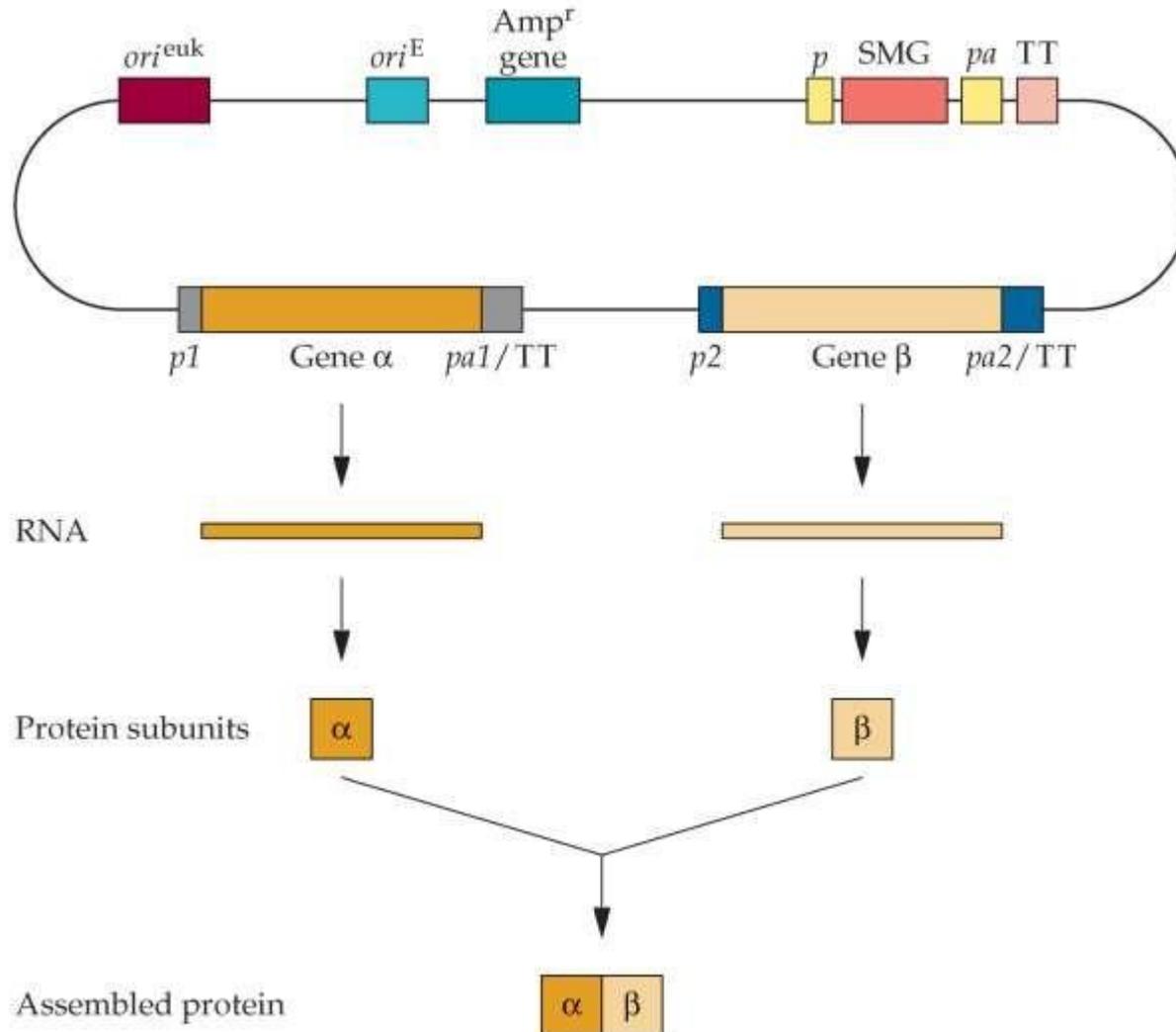
- For the best results, a gene of interest must be equipped with translation control sequences.
- A gene of interest can be fitted with various sequences that enhance translation and facilitate both secretion and purification.
- A Kozak sequence, specific sequence surrounding the AUG start codon, signal sequence, protein affinity tag for purification, proteolytic cleavage site, and stop codon.
- The 5' and 3' UTR increase the efficiency of translation and contribute to mRNA stability.



# Two-Vector Expression System

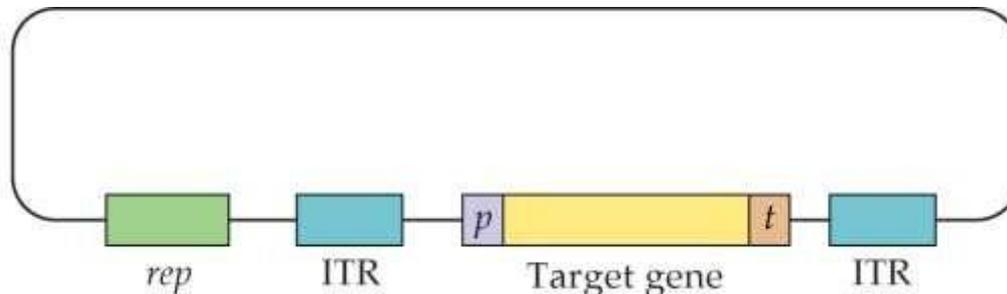


# Two-Vector Expression System



# Baculovirus Vector in Mammalian Cells

- It is possible to use some of the baculovirus vector to express target proteins in mammalian cells
- Because baculovirus cannot replicate in mammalian cells and the polyhedron-deficient strains employed as vectors cannot infect insects. It is a safe system.
- For stable long-term expression, the target gene is inserted between sequences for adeno-associated virus inverted terminal repeat to facilitate the integration into the host cells.



# Selectable Markers for mammalian Expression Vectors

**TABLE 7.7** Selective marker gene systems for mammalian cells

Selective agent	Action of selective agent	Marker gene	Action of marker gene protein
Xyl-A	Damages DNA	Adenine deaminase ( <i>ada</i> )	Deaminates Xyl-A
Blasticidin S	Inhibits protein synthesis	Blasticidin S deaminases ( <i>Bsr</i> , <i>BSD</i> )	Deaminates blasticidin S
Bleomycin	Breaks DNA strands	Bleomycin-binding protein ( <i>Ble</i> )	Binds to bleomycin
G-418 (Geneticin)	Inhibits protein synthesis	Neomycin phosphotransferase ( <i>neo</i> )	Phosphorylates G-418
Histidinol	Produces cytotoxic effects	Histidinol dehydrogenase ( <i>hisD</i> )	Oxidizes histidinol to histidine
Hygromycin B	Inhibits protein synthesis	Hygromycin B phosphotransferase ( <i>Hph</i> )	Phosphorylates hygromycin B
MSX	Inhibits glutamine synthesis	Glutamine synthetase ( <i>GS</i> )	Cells that produce excess glutamine synthetase survive.
MTX	Inhibits DNA synthesis	Dihydrofolate reductase ( <i>dhfr</i> )	Cells that produce excess dihydrofolate reductase survive.
PALA	Inhibits purine synthesis	Cytosine deaminase ( <i>codA</i> )	Lowers cytosine levels in the medium by converting cytosine to uracil
Puromycin	Inhibits protein synthesis	Puromycin <i>N</i> -acetyltransferase ( <i>Pac</i> )	Acetylates puromycin

MSX, methionine sulfoximine; MTX, methotrexate; PALA, *N*-(phosphoacetyl)-*L*-aspartate; Xyl-A, 9- $\beta$ -D-xylofuranosyl adenine.



Thank You